

Sandboxed User Environments with Jupyterhub and Docker

Shreyas Cholia

Donald Winston

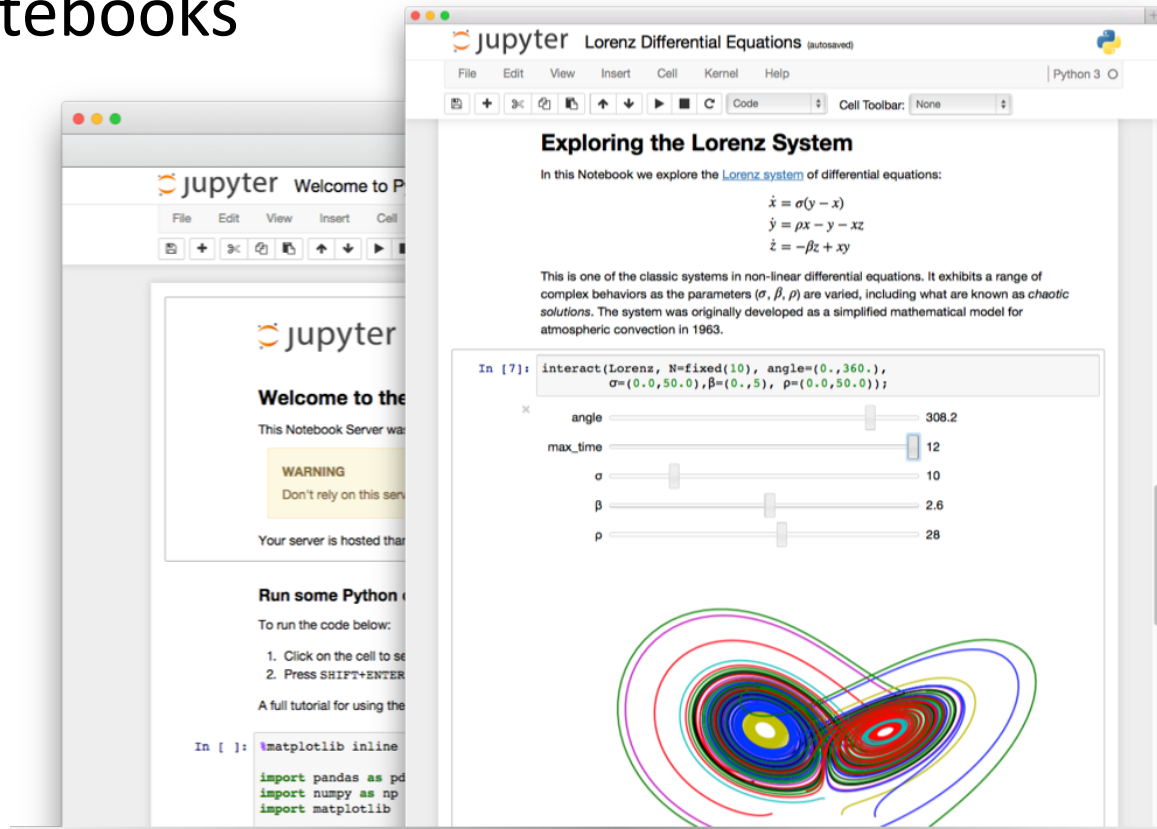
Daniel Gunter

Lawrence Berkeley National Laboratory

Gateways 2016, San Diego

Jupyter

- Interactive notebook environment that can combine code with visualizations, documentation and programmable widgets
- Formerly iPython Notebooks



But Wait .. There's More

- Jupyter also includes
 - File Browser
 - Text Editor (with full syntax highlighting)
 - Terminal

```
jupyter potcar_setup.py ✓ 09/05/2014
File Edit View Language

1 #!/global/ul/s/shreyas/dev/testcif/bin/python2.7
2
3 import os
4 import glob
5 import shutil
6 import subprocess
7
8 pspdir = ""
9
10 count = 0
11 while not os.path.exists(pspdir):
12     if count != 0:
13         print("Invalid vasp dir!")
14     pspdir = raw_input("Please enter full path where the POT_GGA_PAW_PBE, "
15                       "etc. subdirs are present. If you obtained the PSPs "
16                       "directly from VASP, this should typically be the "
17                       "directory that you untar the files to : ")
18
19     print
20
21 targetdir = raw_input("Please enter the fullpath of the where you want to "
22                      "create your pymatgen resources directory: ")
23
24 print
25
26 os.makedirs(targetdir)
27 print("Generating pymatgen resources directory")
28
29 for (parent, subdirs, files) in os.walk(pspdir):
30     for subdir in subdirs:
31         filenames = glob.glob(os.path.join(parent, subdir, "POTCAR*"))
32         if len(filenames) > 0:
33             basedir = os.path.join(targetdir, os.path.basename(parent))
34             if not os.path.exists(basedir):
35                 os.makedirs(basedir)
36             fname = filenames[0]
37             dest = os.path.join(basedir, os.path.basename(fname))
```

jupyter

```
bash-4.1$ ls -l tmp
total 16416
drwxrwxr-x 2 shreyas shreyas 512 Jan 29 2016 build
drwxr-xr-x 2 shreyas shreyas 512 Oct 7 19:13 conf
drwxr-xr-x 3 shreyas shreyas 4096 Oct 7 19:14 conf.d
drwxrwxr-x 4 shreyas shreyas 4096 May 9 2013 dist.eugridpma.info
drwxrwxr-x 5 shreyas shreyas 512 Oct 21 16:51 foo
drwxrwxr-x 5 shreyas shreyas 512 Oct 21 23:01 foobar
-rw----- 1 shreyas shreyas 16384 Apr 30 2015 globus-user-env.sh.swp
drwxrwxr-x 5 shreyas shreyas 512 Oct 14 22:30 msq
drwxrwxr-x 5 shreyas shreyas 512 Jun 3 22:25 pmg
drwxrwxr-x 6 shreyas shreyas 512 Jan 29 2016 pydap-3.1.1
drwxrwxr-x 5 shreyas shreyas 512 Jan 29 2016 pydap-3.2
drwxr-x--- 18 shreyas shreyas 4096 May 26 2015 Python-2.7.10
-rw-r----- 1 shreyas shreyas 16768806 May 23 2015 Python-2.7.10.tgz
drwxr-x--- 5 shreyas shreyas 512 May 26 2015 req
bash-4.1$
```

Files Running Clusters

Select items to perform actions on them.

<input type="checkbox"/>	<input type="text" value="global / homes / s / shreyas / dev"/>
<input type="checkbox"/>	..
<input type="checkbox"/>	bdccclient-2.2.1
<input type="checkbox"/>	bdctest
<input type="checkbox"/>	composetest
<input type="checkbox"/>	data_docker
<input type="checkbox"/>	db_scripts
<input type="checkbox"/>	gridmap-verify
<input type="checkbox"/>	ldap
<input type="checkbox"/>	osg
<input type="checkbox"/>	osg-pki-tools.orig
<input type="checkbox"/>	testcif
<input type="checkbox"/>	tmpvenv
<input type="checkbox"/>	Untitled.ipynb
<input type="checkbox"/>	osg-pki-tools-1.2.11-1.osg32.el5.noarch.rpm
<input type="checkbox"/>	osg-pki-tools.tar.gz

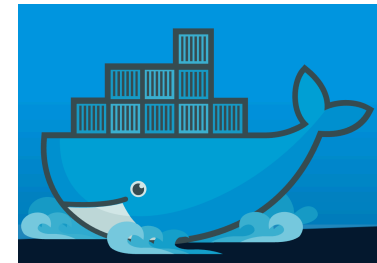
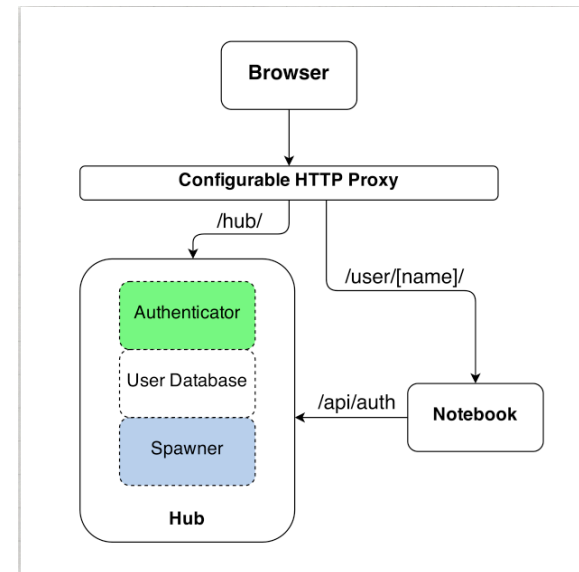
Self Contained environment

- Jupyter can essentially be a fully self-contained interface for computation directly accessible through a web browser
- Fernando Perez has an exciting keynote tomorrow on the possibilities of Jupyter



Complementary Technologies

- Jupyterhub
 - Multiuser Jupyter service for launching notebooks
- Docker
 - Linux containers for creating private user sandboxed “OS” environments
 - You see your own private version of OS, filesystem, PIDs, users etc. i.e. virtualization-lite



Motivating Use Cases

- ***Materials Project: Jupyter for Workshops***

- Containers with the entire materials project software stack pre-installed.
- Tutorial Notebooks
- Local Services (MongoDB and web services)



- ***IDAES: Jupyterhub for Persistent Sandboxes***

- Capture complex software environment in a container
- Enable a development playground with worked examples for hacking on
- Serves as “live documentation”



IDAES

Institute for the Design of Advanced Energy Systems

Goal

- Create a service that allows users to **login** and deploy their own **private, pre-installed Jupyter notebook** environments using **containers**.

Demo Time

We're going to

- Set up Jupyterhub
- Enable user authentication with Github OAuth
- Deploy notebooks in a per user Docker sandbox
- Run additional services in the sandbox
- Live notebook examples

Demo Links

If you want to try this at home

- Jupyterhub setup:

<https://github.com/materialsproject/mp-jupyterhub/>

- Docker Image:

<https://hub.docker.com/r/materialsproject/jupyterhub-singleuser/>

- Docker image build:

<https://github.com/materialsproject/mp-jupyter-docker/>

We're Hiring

- Postdoc to work on Jupyter for human-in-the-loop HPC workflows
- Various CSE positions
- <http://jobs.lbl.gov>

Contact: scholia@lbl.gov